CALIBRATION OF ENSEMBLE SPREAD USING FORECAST SPECTRA

Joshua P. Hacker and David P. Baumhefner

hacker@ucar.edu

National Center for Atmospheric Research

Outline

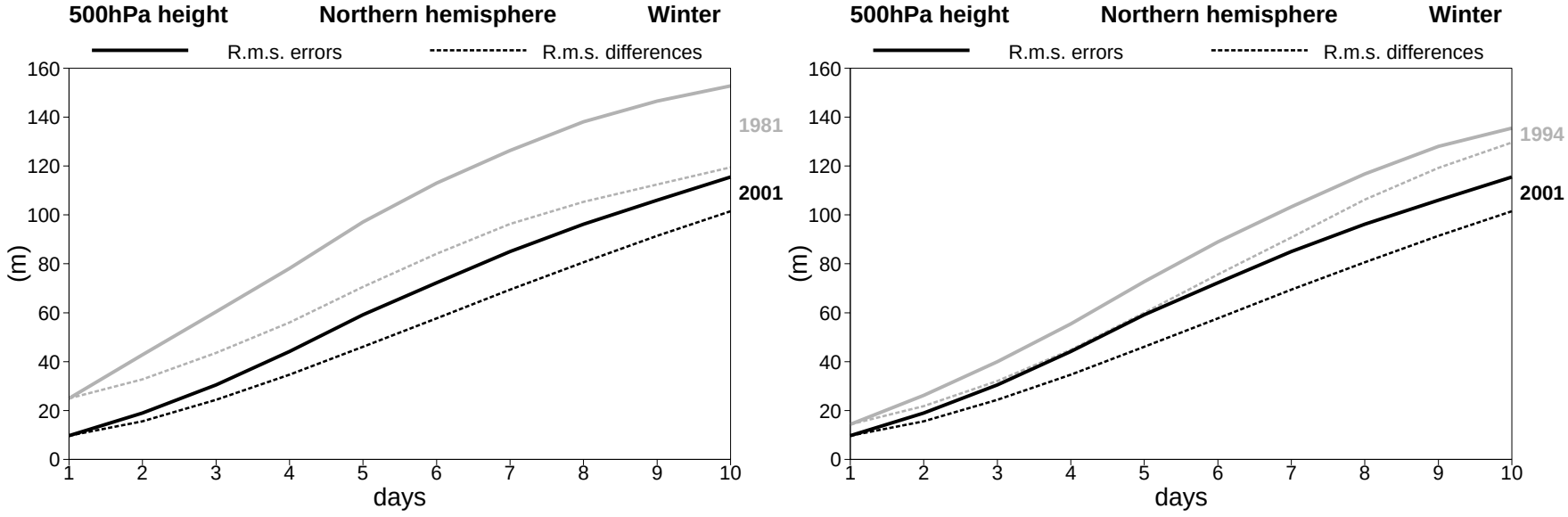
- Motivation
- Analytic response to damping
- Ensemble spread in two similar models
- Ensemble spread in two very different models
- Limitations and complications
- Conclusions

Under-dispersive ensembles

- Overprediction of forecast skill by unknown amount
- Poor prior error-covariance estimates for data assimilation

Under-dispersive ensembles

- Overprediction of forecast skill by unknown amount
- Poor prior error-covariance estimates for data assimilation



Source: Simmons and Hollingsworth 2001

Spread in spectral space

$$\sigma_S^2 = \sum_{k=1}^{N-1} \hat{S}^2 = \sum_{k=1}^{N-1} \left(\hat{P}^2 + \hat{Q}^2 - 2\hat{P}\hat{Q} \right)$$

- Spectral coefficients denoted $\hat{*}$ are functions of wavenumber k.
- P, Q are two members of an ensemble.
- Implied sum over all possible pairs.

A damped model

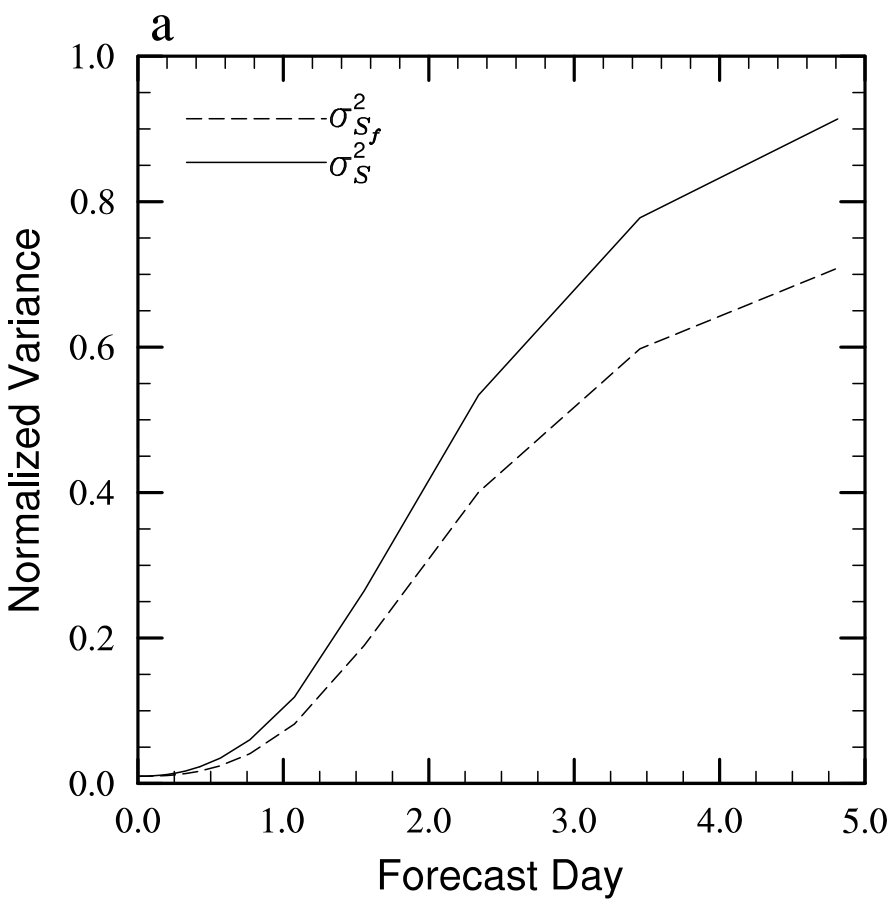
At a single time t and wavenumber k, damp the model with a filter $R(k)$:

$$\hat{p} = R\hat{P}, \quad \hat{q} = R\hat{Q}$$

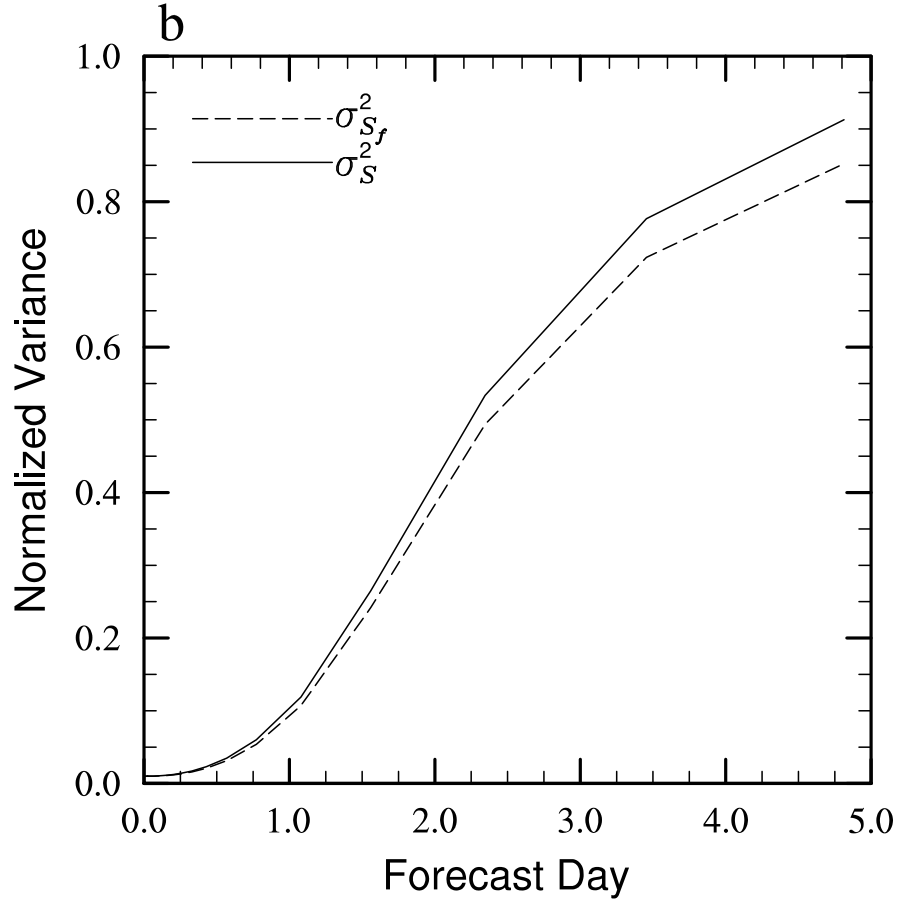
Call it a *deficiency* in scale-dependent spatial variance (forecast amplitude). It results in a damped spread $\hat{S}_f^2$:

$$\begin{aligned} \hat{S}_f^2 &= \hat{p}^2 + \hat{q}^2 - 2\hat{p}\hat{q} \\ &= R^2\left(\hat{P}^2 + \hat{Q}^2 - 2\hat{P}\hat{Q}\right) \\ &= R^2\hat{S}^2 \end{aligned}$$

Consequences



Low resolution



High resolution

How about a real model?

- Error growth depends on case and forecast time
- R is not smooth
- For many forecast systems, we can estimate R

How about a real model?

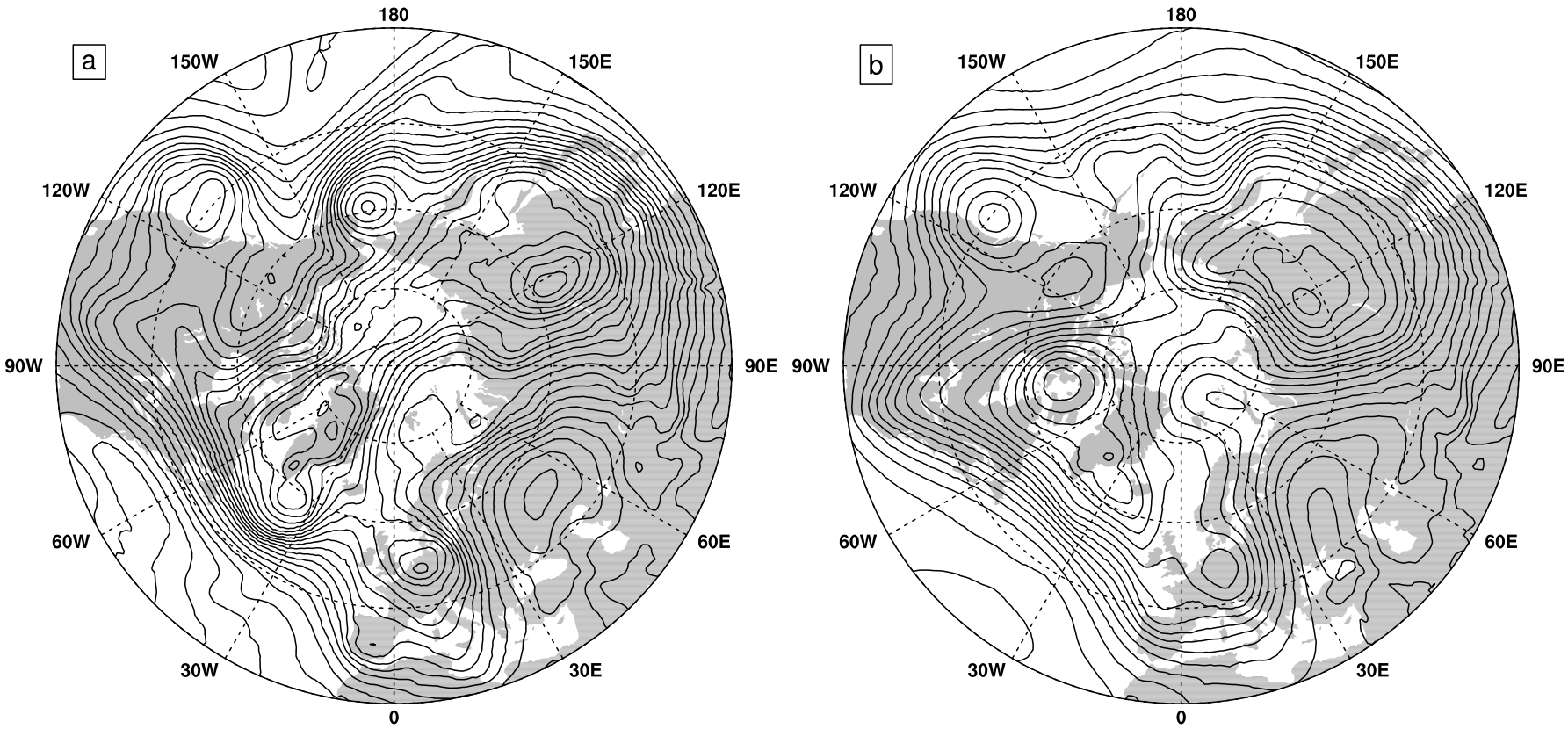
- Error growth depends on case and forecast time
- R is not smooth
- For many forecast systems, we can estimate R

$$
\begin{aligned}
R &= \frac{\hat{p}}{\hat{P}} \\
\hat{S}^2 &= \frac{\hat{S}_f^2}{R^2} \\
\sigma_S^2 &= \sum_{k=1}^{N-1} \hat{S}^2
\end{aligned}
$$

We can calibrate ensemble spread to any reference $\hat{P}$.

Two similar models

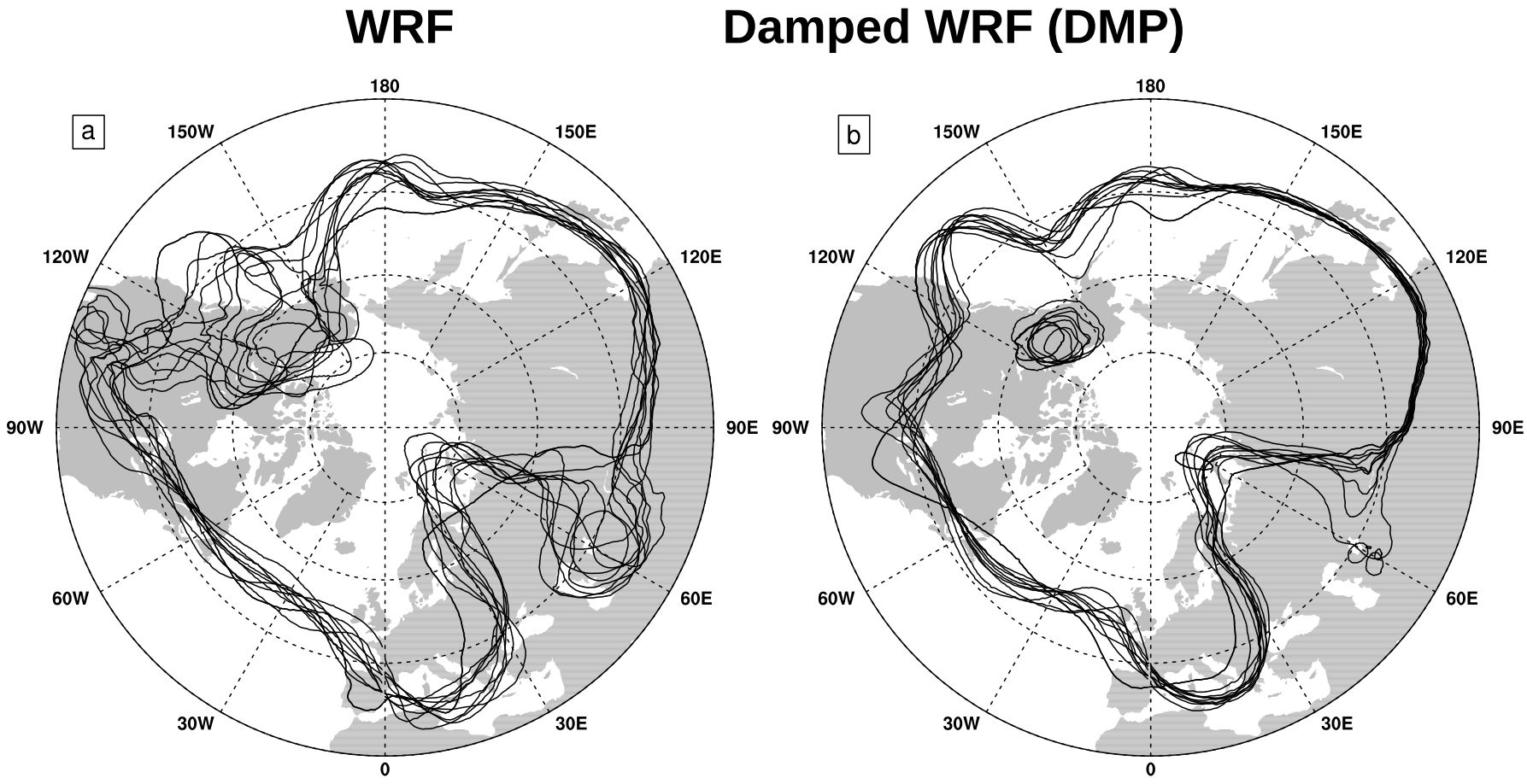
WRF **Damped WRF (DMP)**



Day 6 50.0 kPa geopotential height

WRF-CCM example

Two similar models



Day 6 50.0 kPa geopotential height (5480 m contour)

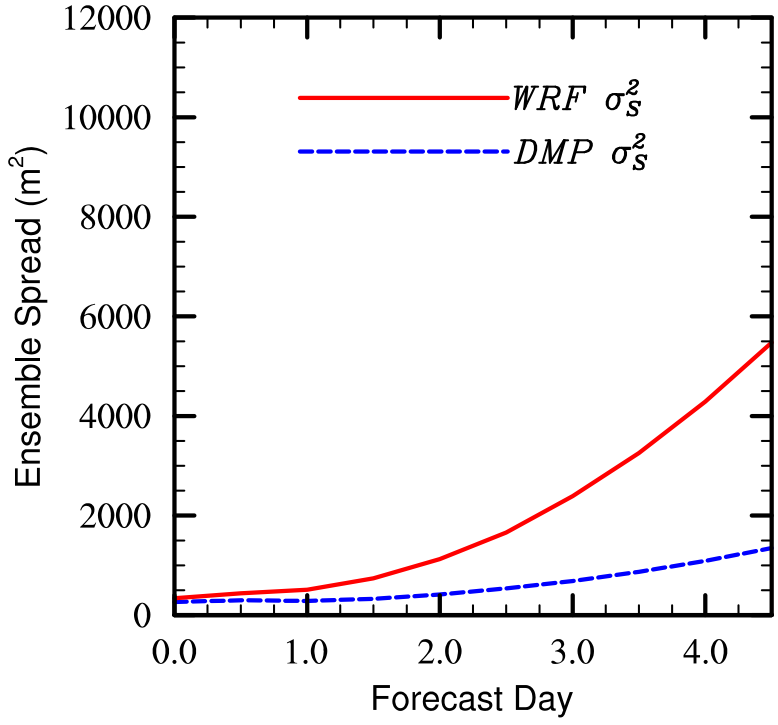
WRF-CCM example

Predicted error growth

Averaged over 6 cases 2001/02 cool season, using 10-member ensembles:

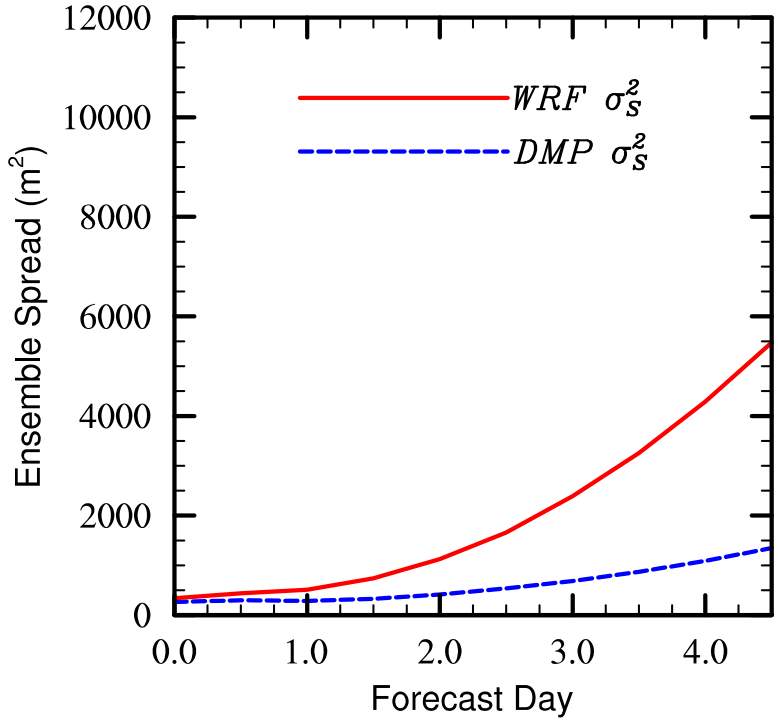
Predicted error growth

Averaged over 6 cases 2001/02 cool season, using 10-member ensembles:



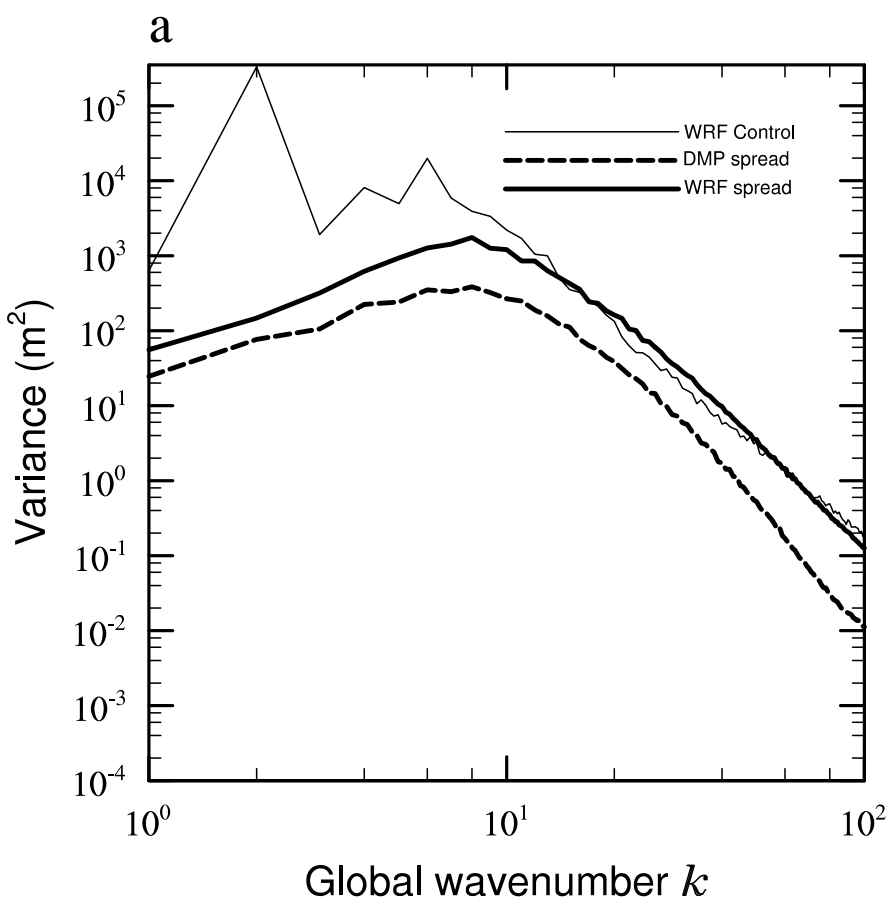
Predicted error growth

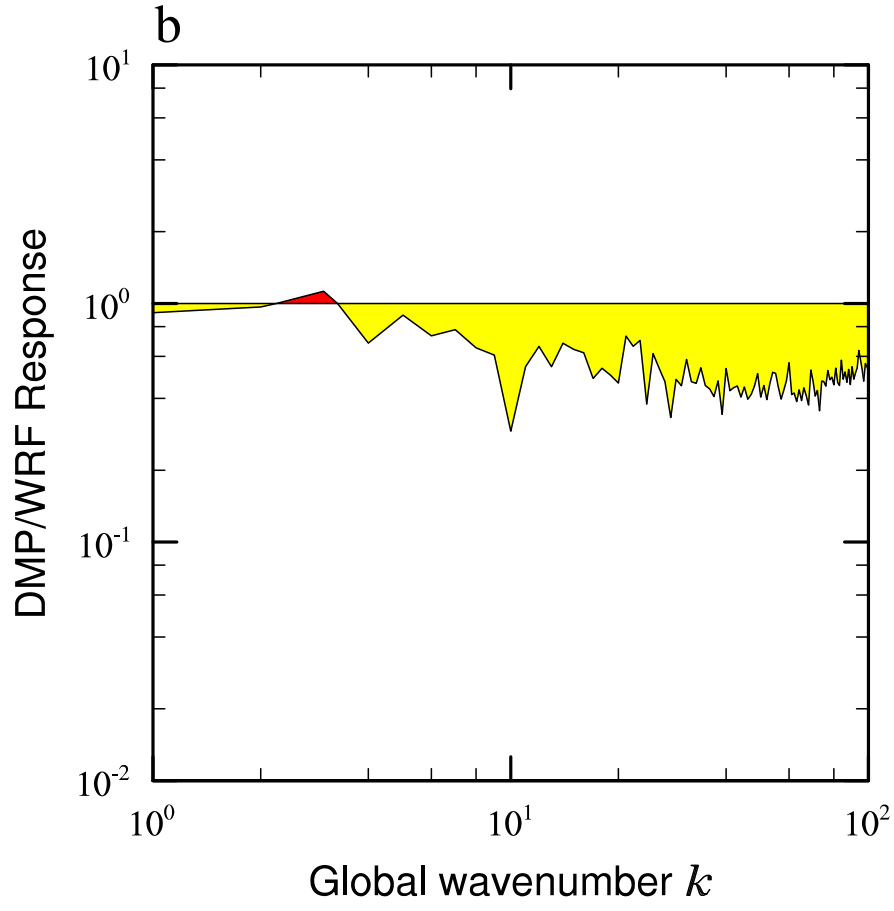
Averaged over 6 cases 2001/02 cool season, using 10-member ensembles:



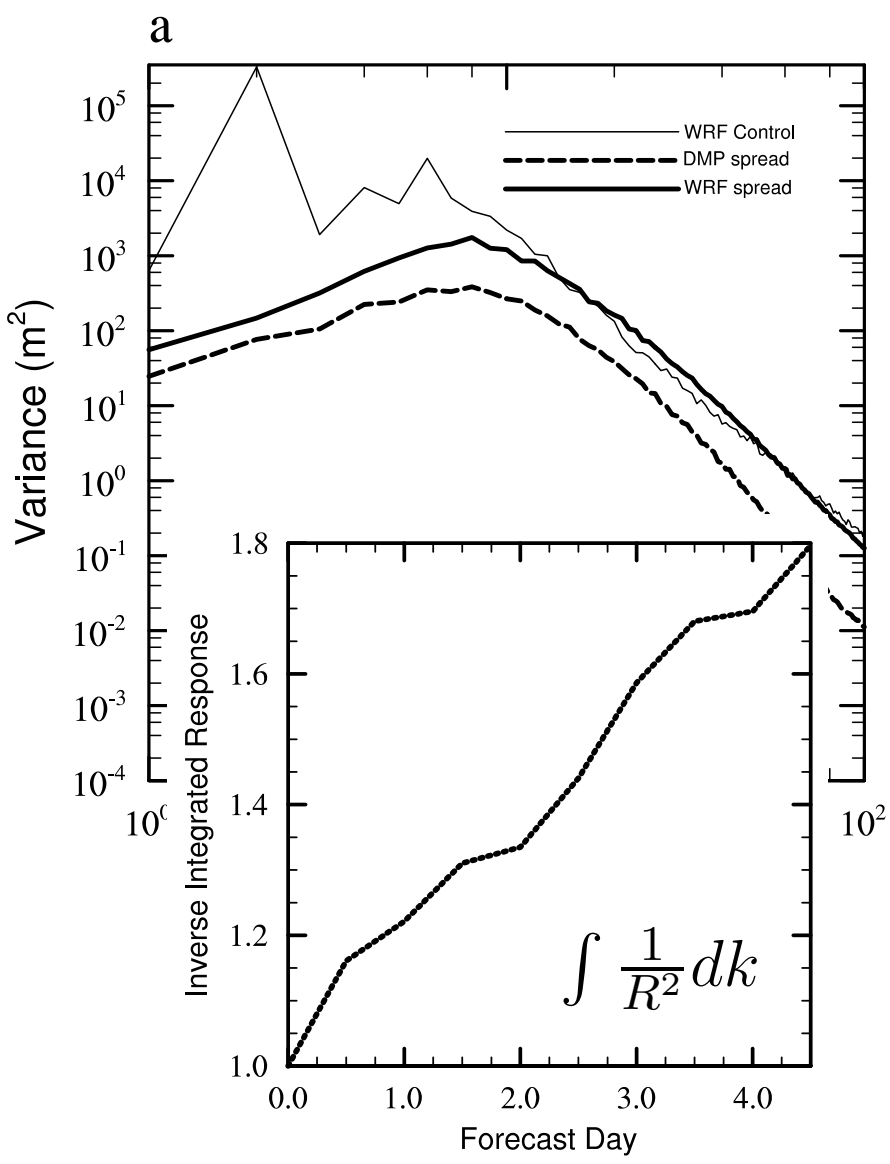
Can we correct this with an estimate of R?

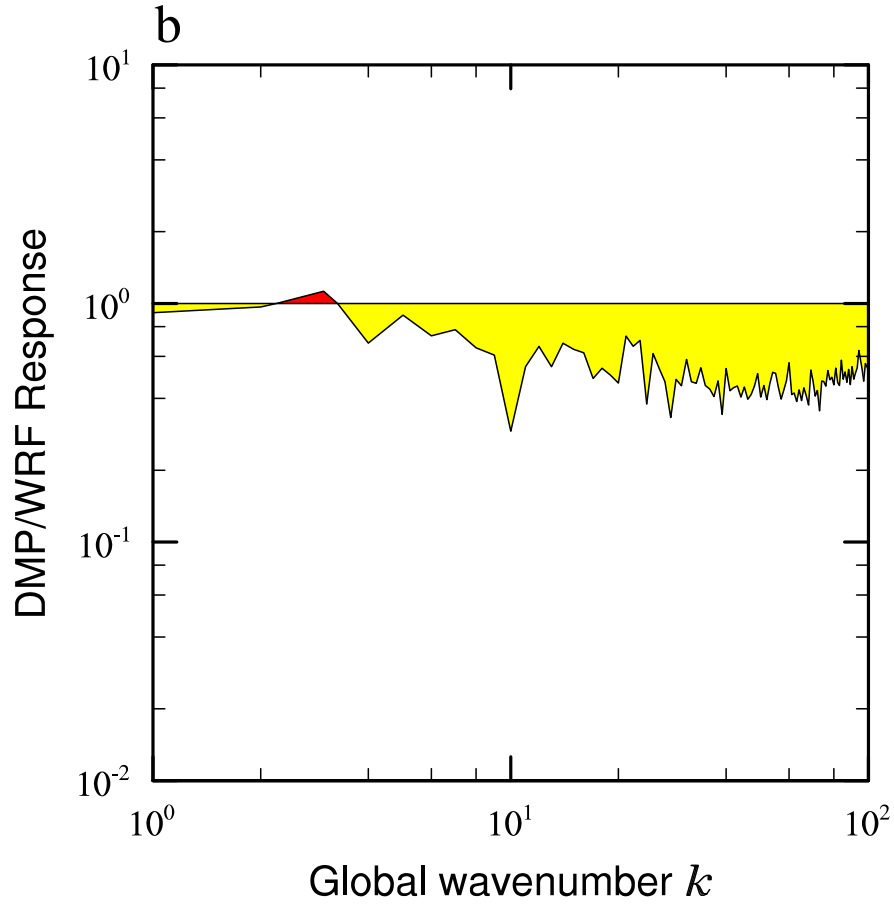
Estimating R



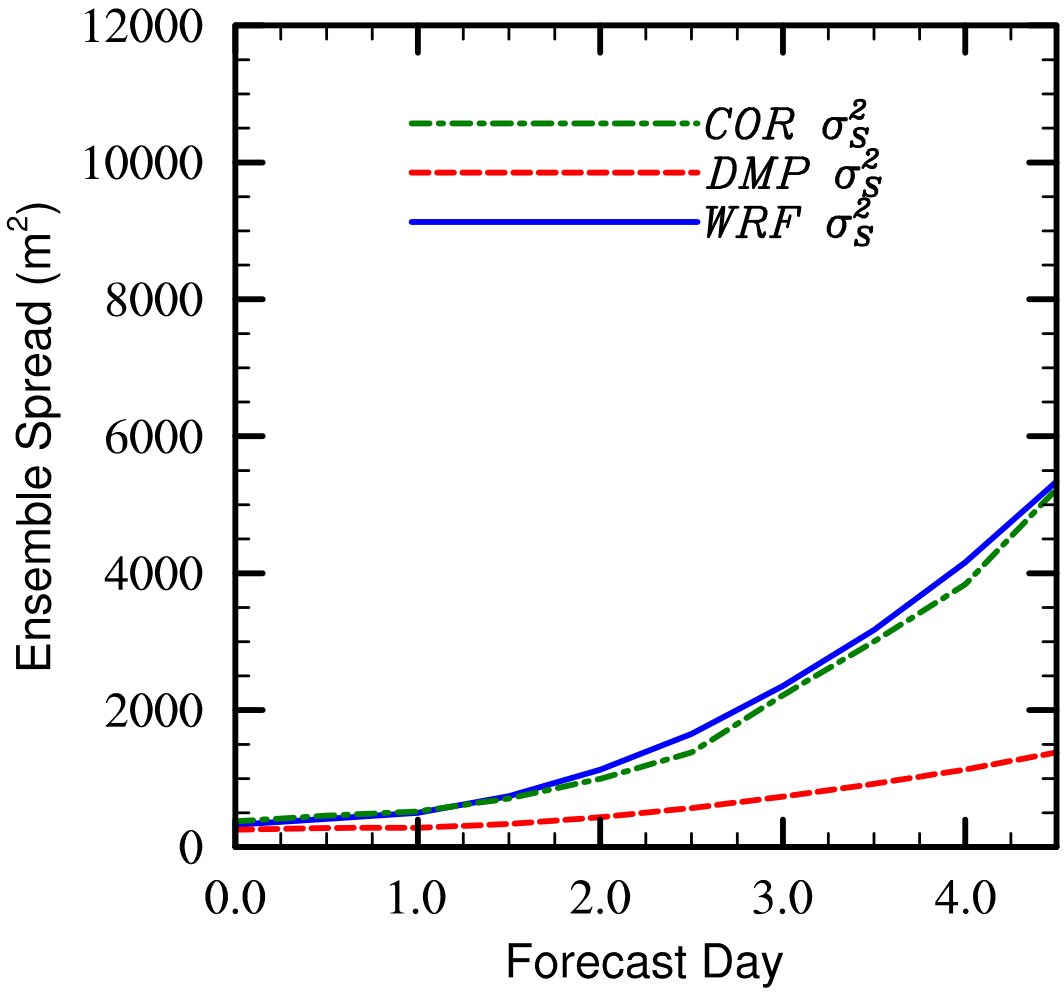


Estimating R

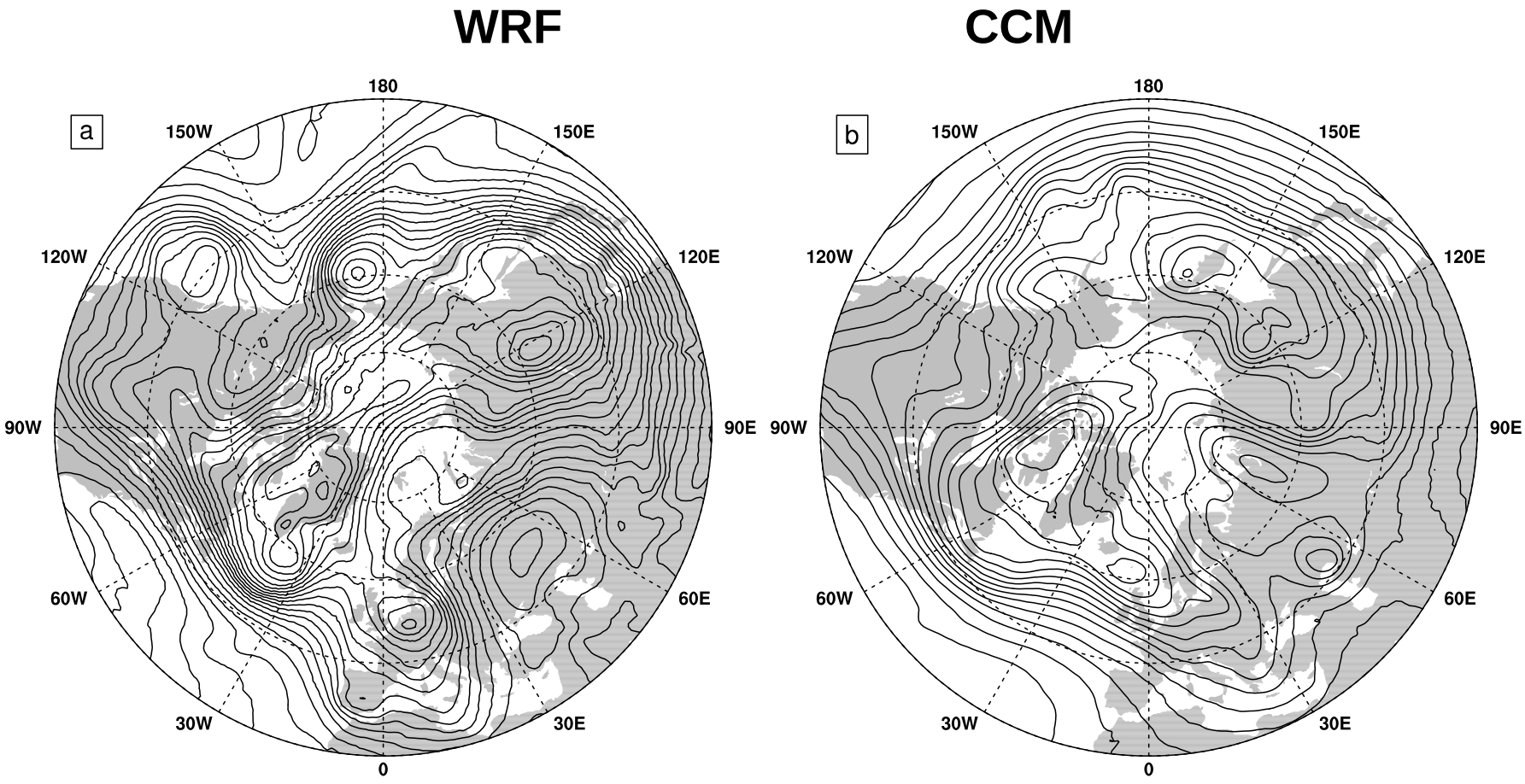




Calibration



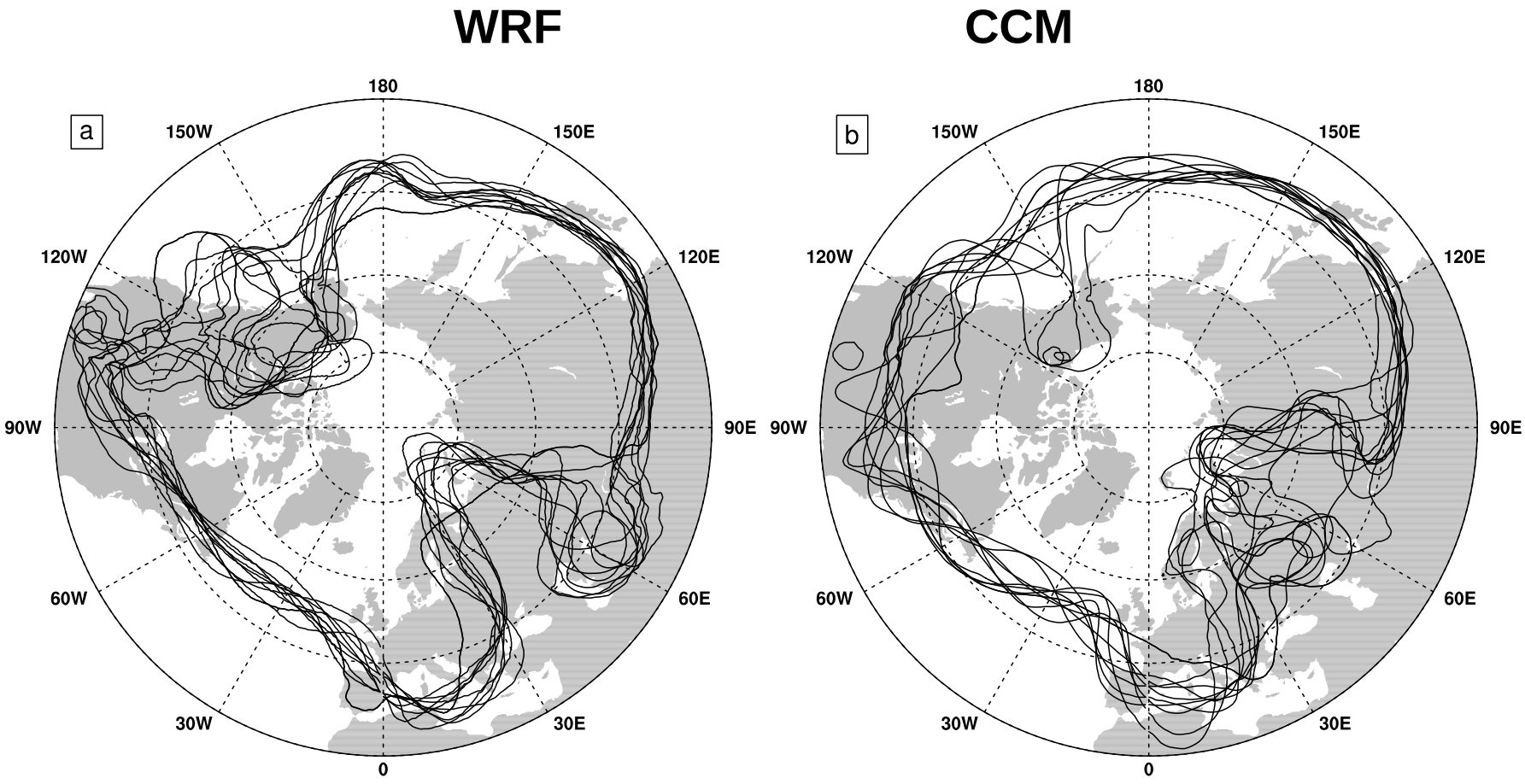
Two different models



Day 6 50.0 kPa geopotential height

WRF-DMP example

Two different models



Day 6 50.0 kPa geopotential height (5480 m contour)

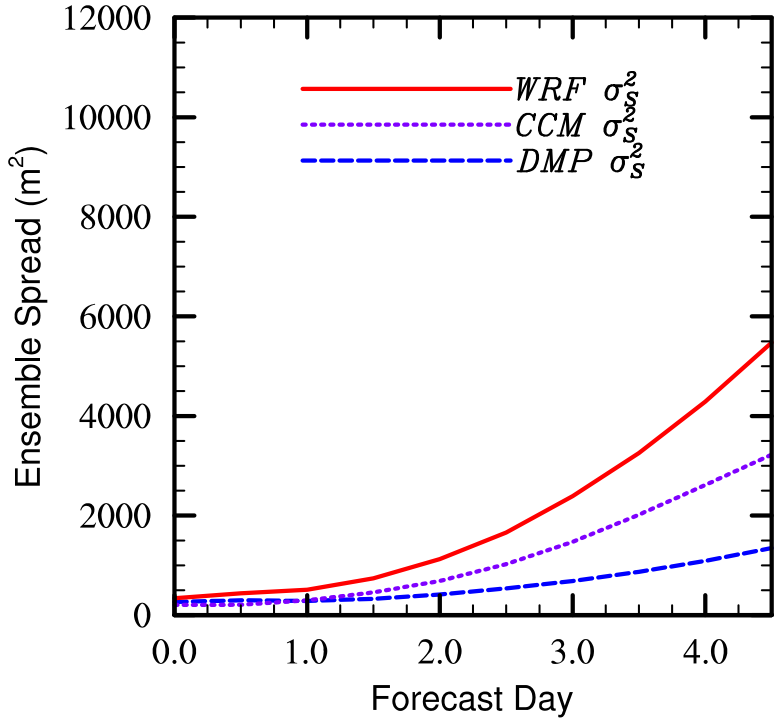
WRF-DMP example

Predicted error growth

Averaged over 6 cases 2001/02 cool season, using 10-member ensembles:

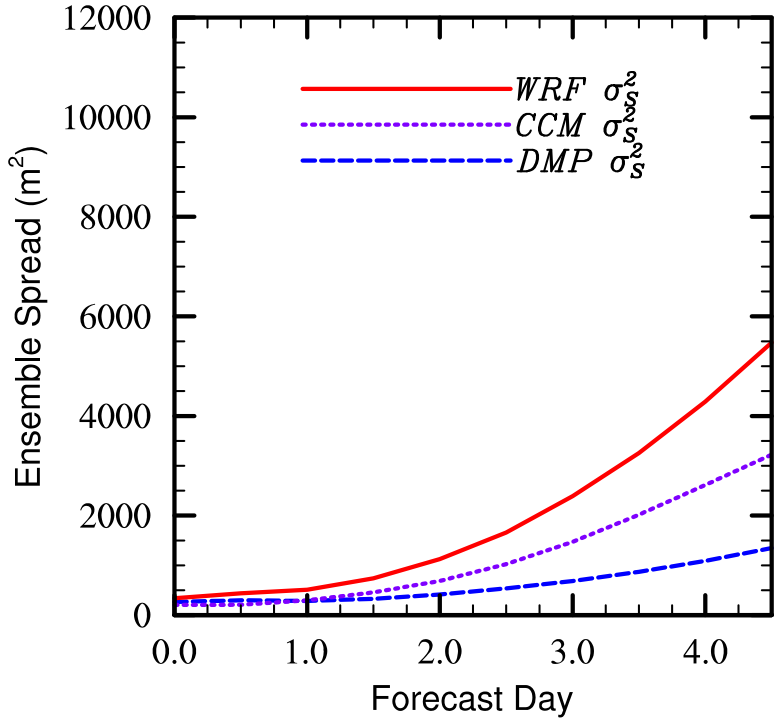
Predicted error growth

Averaged over 6 cases 2001/02 cool season, using 10-member ensembles:



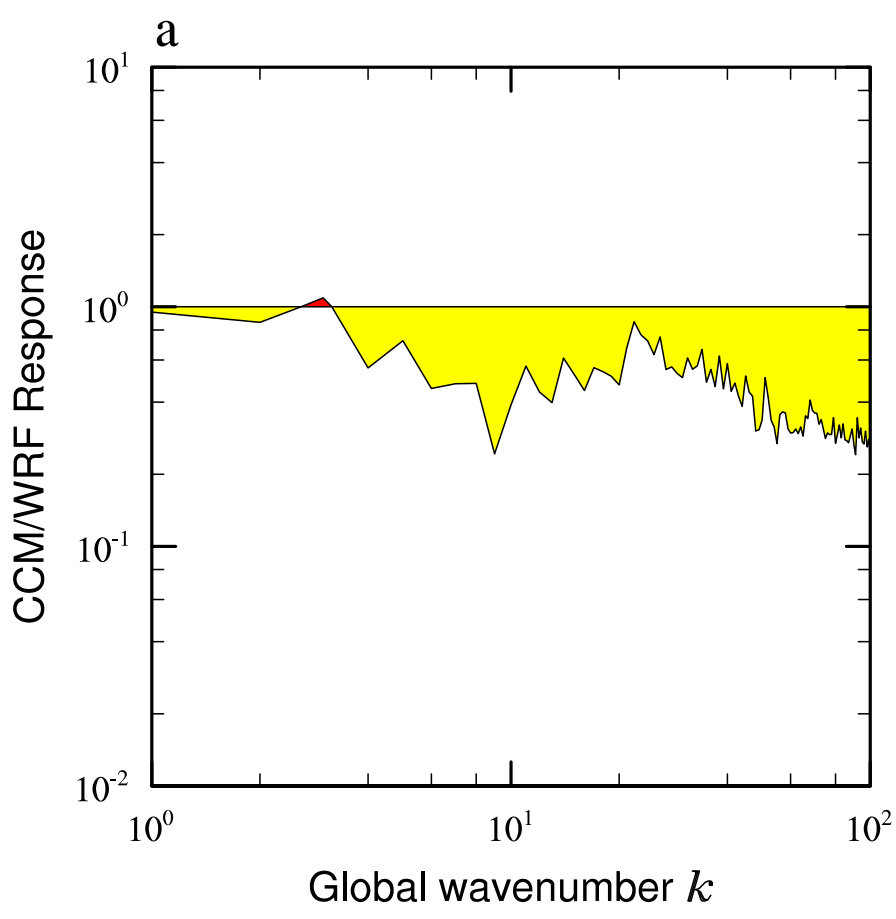
Predicted error growth

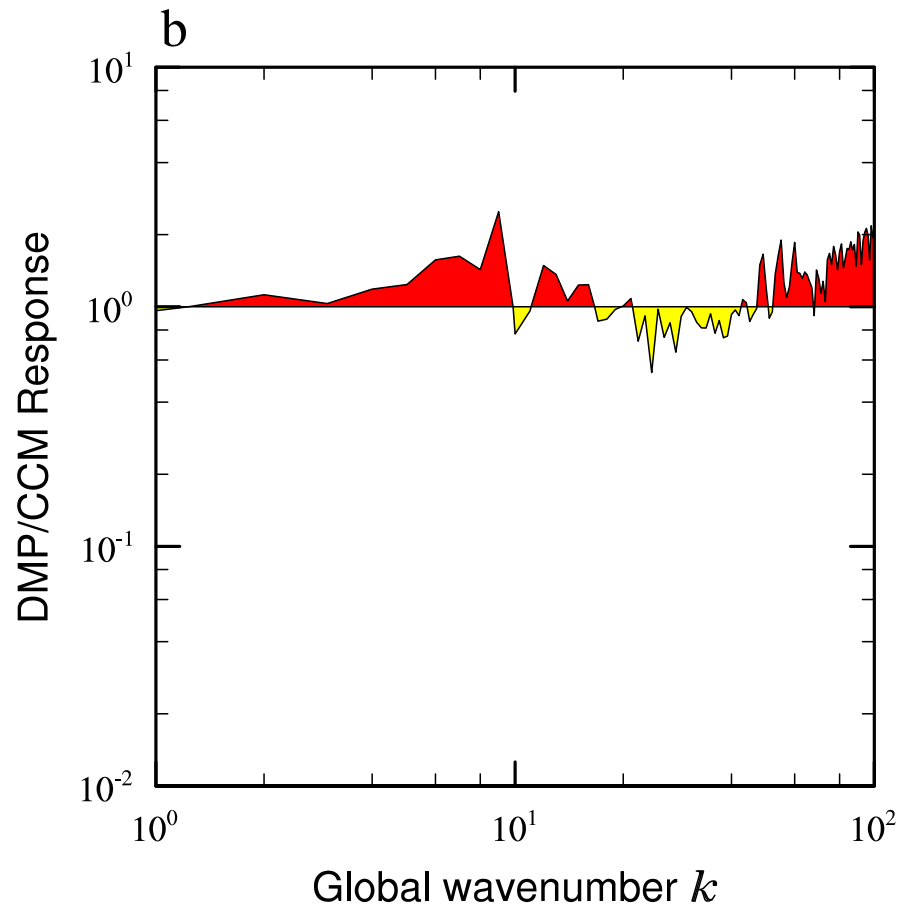
Averaged over 6 cases 2001/02 cool season, using 10-member ensembles:



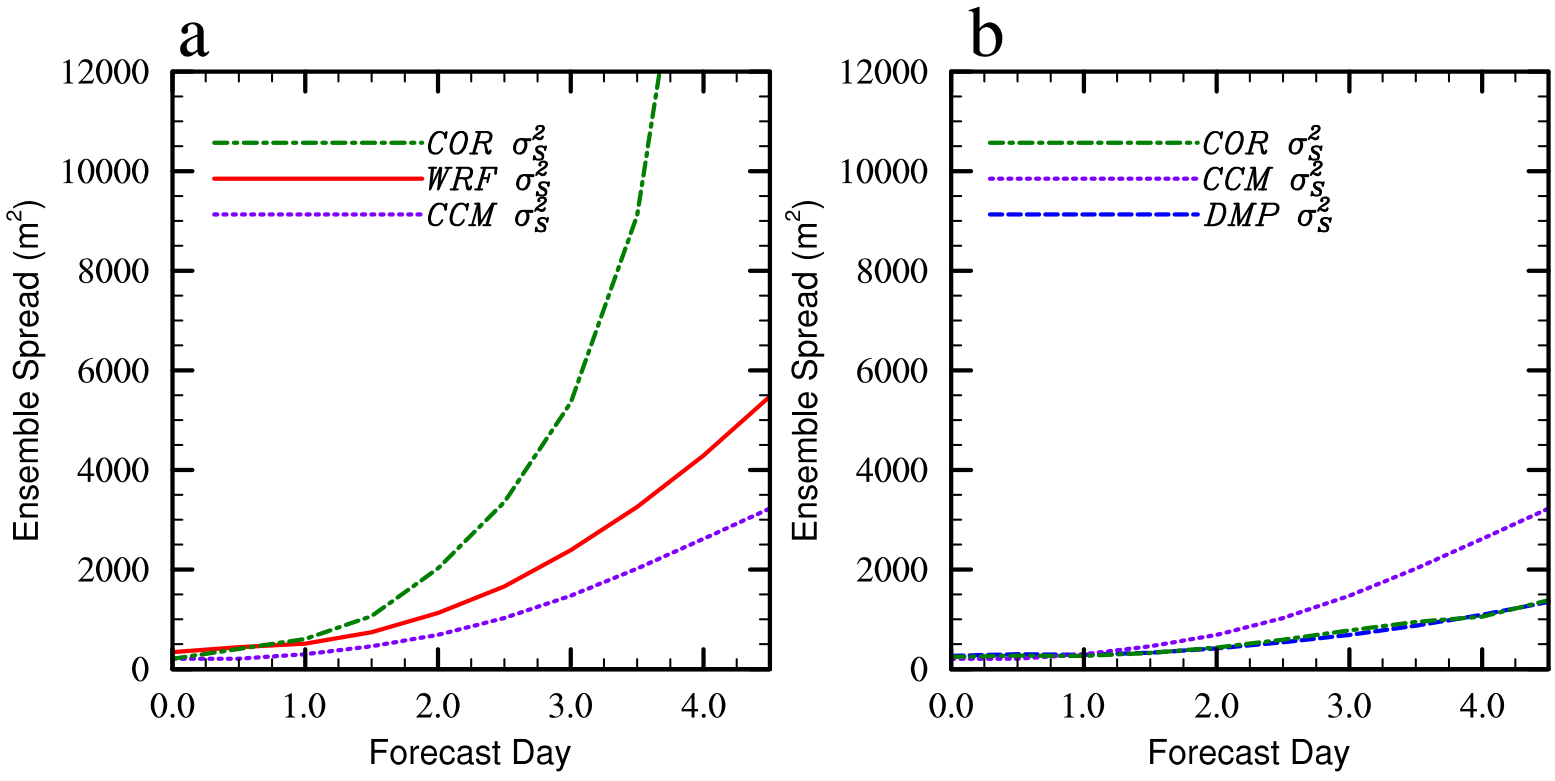
Can we correct this with an estimate of R?

Estimating R

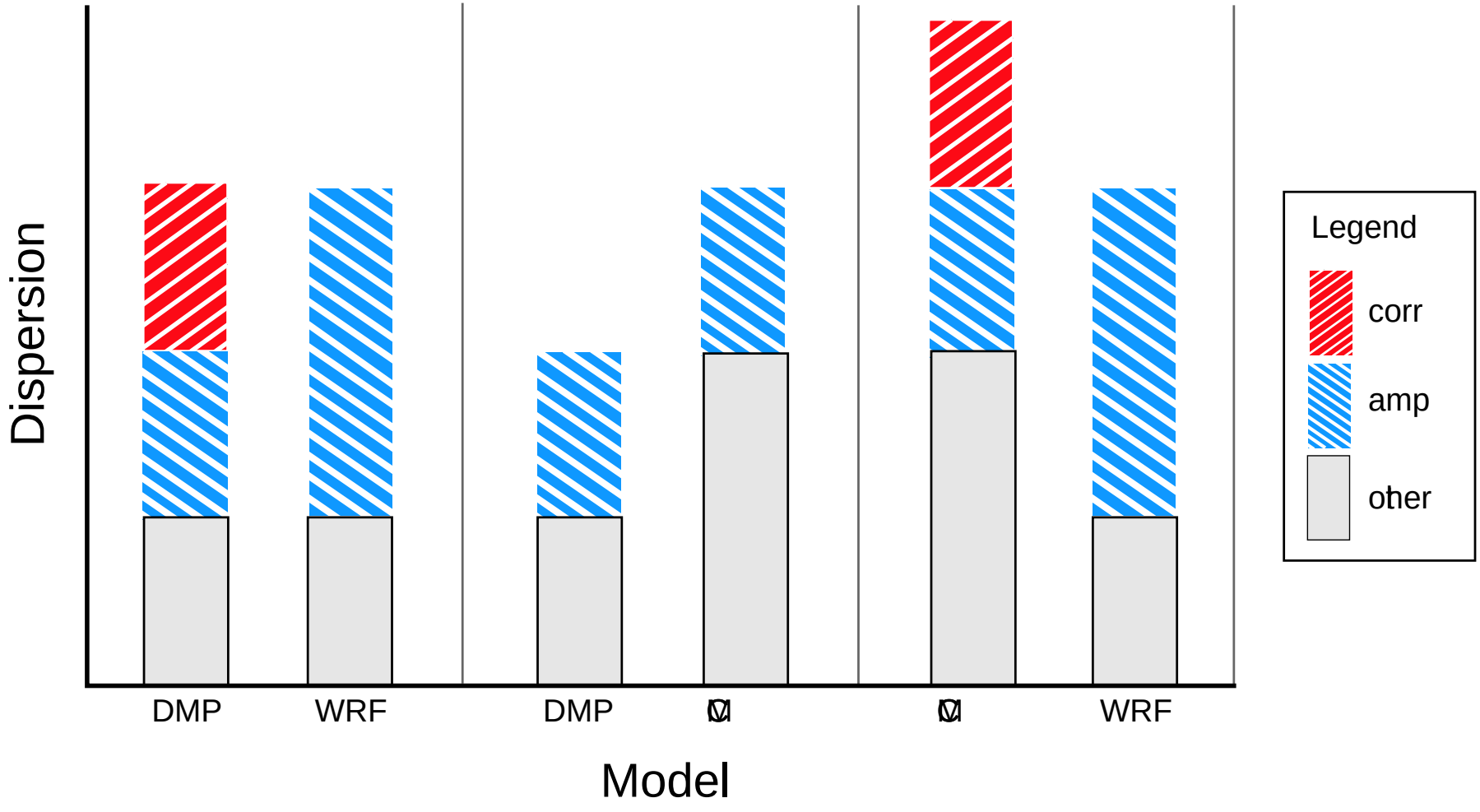




Calibration



Error diagnosis



Complications and limitations

- The calibration only addresses amplitude deficiencies in a model.

- Computing spectra on limited-area domains presents its own challenges.

Conclusions

- Damping in a model will lead to underdispersive ensembles and overly optimistic estimates of predictability.
- Amplitude deficiencies can be corrected with an empirical estimate of the time-dependent ratios between the spectra of different model solutions.
- Residuals (uncorrected spread) provide a measure of the effects of additional model error on ensemble spread.

To do:

- Seek calibration that corrects for a larger class of error.